WAVELET-BASED ESTIMATION OF LONG-MEMORY NOISE IN DIFFUSE OPTICAL IMAGING

C. Matteau-Pelletier, M. Dehaes, F. Lesage

École Polytechnique de Montréal,
Département de génie électrique,
C.P. 6079, succ. centre-ville,
Montréal (Qu.), H3C 3A7, Canada

J.M. Lina

École de technologie supérieure,
Département de génie électrique,
1100, rue Notre-Dame Ouest,
Montréal (Qu.), H3C 1K3, Canada

ABSTRACT

In this work we extend a previously proposed method for functional Magnetic Resonance Imaging (fMRI) to estimate the parameters of a linear model of Diffuse Optical Imaging (DOI) time series. The regression is performed in the wavelet domain to infer drift coefficients at different scales and to estimate the strength of the Hemodynamic Response Function (HRF). This multiresolution approach benefits from the whitening property of the Discrete Wavelet Transform (DWT), and we observe an estimation improvement which is then related to a quantitative measure of $1/f$ noise. The performances of the method are evaluated against a standard spline-cosine GLM approach with simulated HRF and real background physiology.

***Index Terms*—** $1/f$ noise, hemodynamic response function, diffuse optical imaging, discrete wavelets.

1. INTRODUCTION

Diffuse optical imaging has emerged as a non-invasive, low-cost method to quantify brain functional activation [1, 2]. This relatively new technique uses near-infrared spectroscopy to detect local changes in oxy-hemoglobin (HbO) and deoxy-hemoglobin (HbR) concentration in biomedical tissue. Interpretation of DOI data relies on a tightly coupled mechanism where increased metabolic demand and associated hemodynamics of cerebral tissue follows neural firing, and therefore changes in optical signal can reflect neurophysiological activity. The near-infrared spectral range allows *in vivo* tissue imaging at depths reaching a few centimeters and with better temporal accuracy than the Blood Oxygen Level Dependent (BOLD) signal obtained with usual fMRI, a drawback being a diminished spatial resolution due to diffusion.

In fMRI data analysis, linear time-invariant models have been widely used to describe the BOLD response to stimuli and a similar approach has been recently considered in diffuse optical imaging [3],[4]. The currently identified sources of the nuisance fluctuations seen in DOI include cardiac pulsation, respiration, intrinsic blood pressure variation and Mayer waves. Moreover, it has been suggested that DOI noise exhibits a $1/f$-like spectral structure [2] due to either imaging system or physiological mechanisms. In a common form, the technique used to recover the functional response is band-pass filtering thus attenuating the low frequency drifts and high frequency physiology. Data is then averaged across trials (block experiments). However, random-event or complex multi-event protocols require other methods such as the General Linear Model (GLM) where drifts are introduced in the model as regressors and modeled. More recently, these analysis methods have been refined in order to include the effects of non stationary physiology on the estimators. For example, blind principal component analysis (PCA) based filtering [5] has been proposed to reduce physiological signal. A quasi-stationary technique is described in Prince et al. [6] by fitting sinusoids amplitudes and phases to model cardiac and respiratory nuisance signal. This work is extended in Diamond et al. [7] where physiological regressors, measured during the experiment, are included in the analysis. This recent work confirms that modeling physiology in the estimation process can be beneficial.

The purpose of this work is to optimize a wavelet-based estimator for activation estimation in DOI data analysis.We compare this method with other techniques and quantify the $1/f$ noise as well as its effect on the underlying estimator.

2. DISCRETE WAVELET TRANSFORM AND WHITENING PROPERTIES

Discrete wavelets are families of basis functions able to sparsely describe signals in the time-frequency plane. In compact notation, the wavelet transform $\mathcal{W}$ at scale J maps a discrete time signal $x(n)$ to the vector $\mathcal{W}\boldsymbol{x}$ given by

$$\mathcal{W}\boldsymbol{x} \overset{\text{def}}{=} \Big[ax_0^J, dx_0^J, dx_1^{J-1}, \ldots, dx_0^j, \ldots, dx_{2^{-j}N-1}^j, \ldots, dx_0^1, \ldots, dx_{2^{-1}N-1}^1\Big]^T. \quad (1)$$

where the coefficients dx_k^j and ax_k^j are the detail and approximation coefficients at scale j and position k respectively. Time series exhibiting a long-range dependence have slowly decaying autocorrelations and their power-law spectrum can be generally written as $S(f)_{|f|\to 0} \sim \sigma^2 c_\gamma / |f|^\gamma$, where c_γ is a dimensionless function of the spectral exponent γ. It has been shown that the wavelet coefficients of a long-range $1/f$ process have a correlation structure ρ whose magnitude decays as [8] :

$$\left|\rho_{j,j'}^{k,k'}\right|_{2^j k - 2^{j'} k' \to \infty} \sim \mathcal{O}\left(\left|2^j k - 2^{j'} k'\right|^{\gamma-1-2p}\right). \quad (2)$$

Therefore, for $-1 < \gamma < 1$, the intercoefficient correlations, both within and between scales can then be neglected for any wavelet with sufficient number of vanishing moments ($p \geq 1$). However, when dealing with finite time series, increasing the regularity of the wavelet may provoke boundary effects and a compromise has to be negotiated in choosing a reasonable degree. In the following, we have used a Daubechies wavelet with four vanishing moments. Daubechies wavelets are the most compactly supported orthogonal wavelet for any number of vanishing moments, hence it mitigates the extent of intercoefficient correlations introduced by periodic boundary correction.

3. GENERAL LINEAR MODEL

Let $x(t)$ be a signal representing the experimental paradigm and defined as the convolution of a canonical HRF $h(t)$ with a protocol $s(t)$, that is $x(t) \stackrel{\text{def}}{=} (h*s)(t)$. The measured time series $y(t)$ representing either oxy- or deoxy-hemoglobin concentration at a given source-detector pair can be modeled as the sum of the response to the stimulus and a large-scale deterministic drift θ

$$y(t) = \beta x(t) + \theta(t) + \nu(t), \quad (3)$$

where β is a scalar representing the strength of the hemodynamic response. The noise model is assumed to be a serially-correlated error ν. The discrete wavelet transform $\mathcal{W}$ is then applied on both sides of (3) to give

$$\mathcal{W}\boldsymbol{y} = \beta\,\mathcal{W}\boldsymbol{x} + \mathcal{W}\boldsymbol{\theta} + \mathcal{W}\boldsymbol{\nu}, \quad (4)$$

where $\mathcal{W}\boldsymbol{x}$ is assumed to be known. The complexity of the drift can be limited by assuming that fine scale coefficients $d\theta_k^j$, $1 \leq j \leq J_0 - 1$ are zero, so that $\boldsymbol{\theta}$ only describes physiological low-frequency components with $d\theta_k^j$, $J_0 \leq j \leq J$. Let $n_0 = 2^{-J_0+1}N$ be the number of coefficients representing θ. The scale J_0 solely characterizes the complexity of the drift, which is described by a combination of large scale wavelets with the scaling function

$$\theta(t) = a\theta_0^J\,\varphi_0^J(t) + \sum_{j=J_0}^{J}\sum_{k=0}^{2^{-j}N-1} d\theta_k^j\,\psi_k^j(t). \quad (5)$$

Following [3], equation (4) can be written as a standard regression model :

$$\mathcal{W}\boldsymbol{y} = A\boldsymbol{\xi} + \mathcal{W}\boldsymbol{\nu}, \quad (6)$$

where the vector of unknown parameters $\boldsymbol{\xi}$ is defined by

$$\boldsymbol{\xi} \stackrel{\text{def}}{=} \left[a\theta_0^J, d\theta_0^J, \ldots, d\theta_{2^{-J_0}N-1}^{J_0}, \beta\right] \quad (7)$$

with the $N \times (n_0 + 1)$ matrix A . The maximum likelihood (ML) estimate of $\boldsymbol{\xi}$ is given by

$$\widehat{\boldsymbol{\xi}}_{\text{ML}} = \left(A^T\Sigma^{-1}A\right)^{-1} A^T\Sigma^{-1}\,\mathcal{W}\boldsymbol{y}, \quad (8)$$

where Σ is the covariance matrix of the noise. The ML estimators provide an alternative approach to estimation of long-memory errors and have been developed in the wavelet domain by several authors (see [9] for review). The algorithm for the combined ML estimation of both matrix Σ and the linear model parameter vector $\boldsymbol{\xi}$ by mean of the spectral exponent γ is discussed below.

3.1. Model Selection

The parameters selection method presented here provides an additional way of improving the efficiency of the estimator as compared to [3]. We first note that the drift $\boldsymbol{\theta}$ is not orthogonal to the estimated response, hence the overlap between the subspaces can deteriorate the performance. One method for studying this intersection of functional spaces is to compute the correlation coefficients between the drift functions and the stimulus function $x(t)$. A proposed filtering method to control the degeneracy is to remove the drift functions that have a strong correlation with the stimulus [4]. Once a correlation threshold value has been chosen, the columns of the design matrix A corresponding to L wavelet atoms to discard from the drift are removed, giving the reduced matrix B of size $N \times (n_0 + 1 - L)$. The vector $\widehat{\boldsymbol{\xi}'}$ of length $(n_0 + 1 - L)$ is obtained by substituting matrix B in (8) such that

$$\widehat{\boldsymbol{\xi}'} = \left(B^T\Sigma^{-1}B\right)^{-1} B^T\Sigma^{-1}\,\mathcal{W}\boldsymbol{y}. \quad (9)$$

The estimate $\widehat{\boldsymbol{\xi}}$ is then recovered by zero-padding at all indexes where the drift is assumed to be zero.

The selection of the scale J_0 that characterizes the number of degrees of freedom of the drift also has a profound effect on estimation of β. In this work, we select the values of J_0 and the correlation threshold that minimize the empirical estimate of the mean squared error for $\hat{\beta}$, based on several simulations of the null hypothesis (rest) condition (see section 4.1).

In the wavelet domain, it is assumed that the error terms are not correlated across time because the wavelet coefficients of a $1/f$ process at scale j are a set of stationary independent identically distributed variables with zero mean and S_{d_j} variance [8]. The matrix Σ is thus approximately diagonalized by the DWT and takes the form

$$\Sigma = \text{diag}\left\{S_{a_J}, S_{d_J}, \ldots, (S_{d_2}, \ldots, S_{d_2}), (S_{d_1}, \ldots, S_{d_1})\right\}, \tag{10}$$

where S_{a_J} and S_{d_j}, $j \in [1, 2, \ldots, J]$, are the variances of the approximation and wavelet coefficients. A preliminary hypothesis would be to assume that the variances of the noise at different scales are equal. Thus, the error covariance matrix is the identity matrix and this first estimation scheme is an ordinary least squares (OLS) analysis. However, the variances of the noise at different scales may be quite different. We used the wavelet-generalized least squares (WLS) algorithm developed in [10] to approximate the maximum likelihood estimator of both model and noise parameters in an iterative fashion. Typically the algorithm requires three iterations or less until the change in successive parameter estimates is less than 0.1%.

4. EXPERIMENTS

All DOI data collected from a CW NIR commercial system (Techen CW5) are preprocessed to obtain a concentration measure from raw photon fluence signals. The optical probes (laser sources and detectors) are positioned on the scalp of the subjects and closest source-detector (optode) pairs are kept for analysis. The photon fluence Φ at wavelength $\lambda_i = 690$, 830 nm is then converted to a variation in optical density ΔOD through the Beer-Lambert law

$$\Delta OD(t, \lambda_i) = -\ln\left(\frac{\Phi(t, \lambda_i)}{\Phi_0(t, \lambda_i)}\right). \tag{11}$$

The next step is to translate the ΔOD measures to concentration changes by using the extinction coefficients ε of both chromophores :

$$\begin{bmatrix} \Delta C_{\text{HbO}}(t) \\ \Delta C_{\text{HbR}}(t) \end{bmatrix} \propto \begin{bmatrix} \varepsilon_{\text{HbO}}^{\lambda_1} & \varepsilon_{\text{HbR}}^{\lambda_1} \\ \varepsilon_{\text{HbO}}^{\lambda_2} & \varepsilon_{\text{HbR}}^{\lambda_2} \end{bmatrix}^{-1} \begin{bmatrix} \Delta OD(t, \lambda_1) \\ \Delta OD(t, \lambda_2) \end{bmatrix}. \tag{12}$$

Each change in concentration is computed for selected optode pairs and are modeled in the definition of the general linear model in the previous section.

4.1. Simulated evoked response

In this section, we present the results of simulation studies used to validate wavelet-based estimation by comparison to a GLM approach that modeled the instrument and physiological drifts as a sum of polynomial and sine functions. Realistic simulations are performed by adding optical data measured at rest to a deterministic evoked response of fixed strength β. Baseline DOI data were collected over the motor area of two subjects. Data from 14 source-detector pairs were included in the analysis. For each channel, we computed the estimate $\widehat{\boldsymbol{\xi}}$ with 10 distinct random stimuli using three estimation schemes : spline/cosine GLM, OLS and WLS. The stimulus paradigm followed an event-related design with a uniformly distributed 4 to 20 seconds interstimulus interval over the 6.8 minute runs. Each stimulus time series was composed of alternating segments of 1 and 0. Each segment of 1 lasted for 2 seconds. We used the model of canonical HRF provided by the SPM package with small values $\beta_{\text{HbO}} = 1 \times 10^{-6}$ and $\beta_{\text{HbR}} = -2 \times 10^{-7}$.

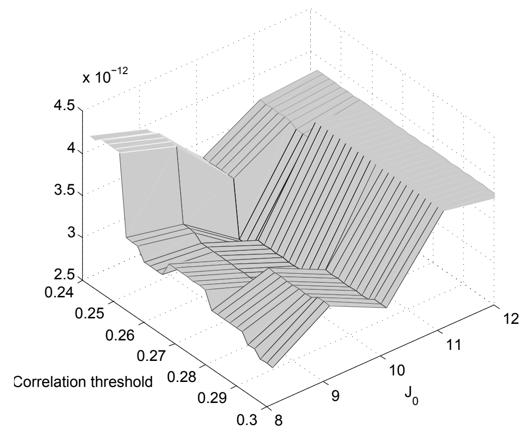

Fig. 1. Mean squared error as a function of the drift finest scale J_0 and the correlation threshold. Note that minimum error is attained for $J_0 = 10$. The gray scale is arbitrary.

The scale J_0 and the correlation threshold above which drift basis functions are discarded were varied in the case of OLS and WLS to investigate their impact on the performance of those estimators. Efficiency was first quantified using the mean squared error between simulated and estimated value of β. Fig. 1 shows this error for ΔC_{HbO} of one subject as a function of the drift complexity J_0 and the correlation threshold in the region of minimal error. The efficiency of the method varies greatly depending on the number of degrees of freedom $n_0 = 2^{-J_0+1}N$ of the baseline drift, and to a lesser extent to the correlation threshold. In the case of ΔC_{HbO}, the minimum of the criterion is reached for $J_0 = 10$ and a correlation value of 0.25. Similar values were obtained for the deoxy-hemoglobin concentration measurements. The overall performance of all three methods, the general linear model with periodic drifts (COS), Ordinary Least Square (OLS), and Wavelet Least Square (WLS), can be compared in terms of efficiency of estimation of the parameter β. These results are shown in Fig. 2, which shows the mean error and the error standard deviation. It can be seen that WLS is more efficient than the spline/cosine GLM approach in the case of ΔC_{HbO}, but that all methods are comparable for ΔC_{HbR}. The data suggest a different correlation structure between HbO and HbR. In order to validate this, we estimated the average spectral exponent, γ, describing the long memory noise for each signal in Table 1.

The much greater spectral exponent for HbO is likely to

Table 1. Mean and standard deviation of the spectral exponent γ for HbO and HbR

	HbO	HbR
Mean	0.550	0.169
Standard deviation	0.206	0.185

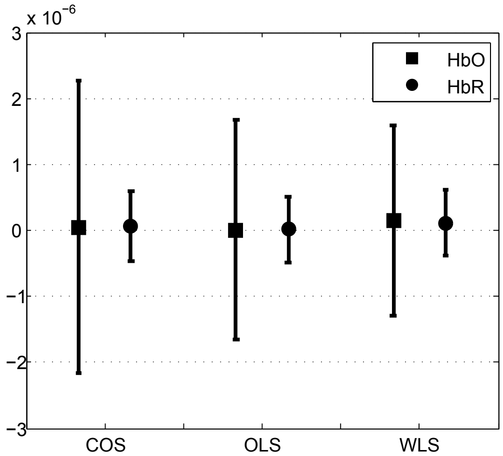

Fig. 2. Standard deviations and biases computed for three models of linear regression : spline/cosine GLM, OLS and WLS. Left : HbO concentration. Right : HbR concentration. Note how the weighted estimation reduces the standard deviation (at the expense of a slightly increased empirical bias illustrating the trade-off between modeling and estimation confidence).

benefit from the whitening effect of the DWT, as opposed to the HbR case where the null value is contained in the standard deviation range of the γ estimate. These empirical findings stand in favor of the presented method as a more efficient estimator for DOI data contamined with $1/f$ noise. The above technique was also evaluated on an event related finger tap experiment with two subjects. Activation was clearly localized over the motor area and results were consistent [11]. Only the above simulations however provide a truth basis to evaluate the technique.

5. CONCLUSION

This work has addressed the problems of characterizing the $1/f$ noise and estimating the strength of hemodynamic response in diffuse optical imaging. Physiological drift were included as part of our model which is distinct from the previous WLS estimator and from a recent work based on continuous wavelet technique [12]. Analysis performed on experimental DOI data provided us with a quantitative measure of the $1/f$ noise for HbO and HbR concentration measurements. These results reflect the observations made on the performance of WLS as compared to another GLM-based estimator.

6. REFERENCES

[1] A. Yodh and B. Chance, "Spectroscopy and imaging with diffusing light," *Phys. Today*, vol. 48, pp. 34–40, 1995.

[2] M. A. Franceschini and D. A. Boas, "Noninvasive measurement of neuronal activity with near-infrared optical imaging," *Neuroimage*, vol. 21, pp. 372–386, 2004.

[3] F. G. Meyer, "Wavelet-based estimation of a semiparametric generalized linear model of fMRI time-series," *IEEE Trans Med Imaging*, vol. 22, no. 3, pp. 315–322, 2003.

[4] J. Cohen-Adad, S. Chapuisat, J. Doyon, J.-M. Lina, H. Benali, and F. Lesage, "Application of the general linear model to response estimation in optical imaging," *Medical Imaging Analysis*, 2007, to appear.

[5] Y. Zhang, D. H. Brooks, and D. A. Boas, "A haemodynamic response function model in spatio-temporal diffuse optical tomography," *Phys. Med. Biol.*, vol. 50, pp. 4625–4644, 2005.

[6] S. Prince, V. Kolehmainen, J. P. Kaipio, M. A. Franceschini, D. Boas, and S. R. Arridge, "Time-series estimation of biological factors in optical diffusion tomography," *Phys. Med. Biol.*, vol. 48, no. 11, pp. 1491–1504, 2003.

[7] S. G. Diamond, T. J. Huppert, V. Kolehmainen, M. A. Franceschini, J. P. Kaipio, S. R. Arridge, and D. A. Boas, "Dynamic physiological modeling for functional diffuse optical tomography," *Neuroimage*, vol. 30, no. 1, pp. 88–201, 2006.

[8] A. H. Tewfik and M. Kim, "Correlation structure of the discrete wavelet coefficients of fractional brownian motion," *IEEE Trans. Inf. Theory*, vol. 38, pp. 904–909, 1992.

[9] G. W. Wornell, *Signal processing with fractals: a wavelet- based approach*, Prentice Hall, 1996.

[10] M. J. Fadili and E. T. Bullmore, "Wavelet-generalized least squares: a new BLU estimator of linear regression models with 1/f errors," *Neuroimage*, vol. 15, pp. 217–232, 2002.

[11] C. Matteau-Pelletier, M. Dehaes, F. Lesage, and J.-M. Lina, "1/f noise in diffuse optical imaging and wavelet-based response estimation," *IEEE Trans Med Imaging*, 2007, submitted.

[12] J.-M. Lina, M. Dehaes, C. Matteau-Pelletier, and F. Lesage, "Complex wavelets applied to diffuse optical spectroscopy for brain activity detection," *Optics Express*, vol. 16, no. 2, pp. 1029–1050, 2008.